

# Économétrie des Séries Temporelles Appliquées à la Finance

Arnaud Buccafurri

L3 Mathématiques mineur Économie à l'Université Paris 8 Vincennes-Saint-Denis

Mars 2025

## Résumé

Dans ce projet, nous nous intéresserons aux séries temporelles. Nous chercherons à les introduire ainsi qu'à étudier leur propriété. Nous aborderons différentes notions tel que les moyennes mobiles, le lissage exponentiel afin d'introduire le modèle ARIMA.  
Par la suite, nous prendrons le cours du CAC40 afin d'estimer les paramètre dans un modèle ARIMA(p,d,q) puis nous terminerons par une prédiction des valeurs futures à l'aide de ce même modèle.



# Table des matières

<b>1</b>	<b>Introduction aux Séries Temporelles</b>	<b>4</b>
1.1	Généralités . . . . .	4
1.2	Estimateur des Moindres Carrés Ordinaires . . . . .	5
1.3	Tendance . . . . .	5
1.4	Saisonnalité . . . . .	6
1.4.1	Modèle trimestriel de Buys-Ballot . . . . .	6
1.5	Bruit Blanc . . . . .	6
<b>2</b>	<b>Moyenne Mobile</b>	<b>7</b>
2.1	Généralité . . . . .	7
2.2	Utilisation des Moyennes Mobiles . . . . .	8
2.3	Valeurs Propres . . . . .	8
2.4	Exemple de Moyennes Mobiles . . . . .	9
2.4.1	Forme générale . . . . .	9
2.4.2	Moyenne Mobile Paire . . . . .	9
2.4.3	Moyenne Mobile Impaire . . . . .	9
2.4.4	Moyenne Mobile Différenciée . . . . .	9
<b>3</b>	<b>Lissage Exponentiel</b>	<b>10</b>
3.1	Lissage Exponentiel Simple . . . . .	10
3.1.1	Ordre 1 . . . . .	10
3.1.2	Itération . . . . .	10
3.2	Lissage Exponentiel Double . . . . .	10
3.2.1	Ordre 1 . . . . .	10
3.2.2	Formule classique . . . . .	11
3.2.3	Itération . . . . .	11
<b>4</b>	<b>Introduction au Modèle ARIMA</b>	<b>12</b>
4.1	Les Espaces $\mathcal{L}^2$ . . . . .	12
4.2	Vecteurs et Processus Gaussiens . . . . .	12
4.2.1	Vecteurs Gaussiens . . . . .	13
4.2.2	Processus Gaussien . . . . .	13
4.3	Régression dans $\mathcal{L}^2$ . . . . .	13
4.3.1	Nombre fini de variables . . . . .	13
4.3.2	Nombre infini de variables . . . . .	13
4.3.3	Opérateur de Projection Linéaire . . . . .	13
4.3.4	Prévision Linéaire . . . . .	14
4.4	Polynômes Opérateur Retard et Avance . . . . .	14
4.5	Les Séries Stationnaires . . . . .	14
4.6	Modèle AR(p) . . . . .	15
4.6.1	AR(1) . . . . .	15
4.7	Modèle MA(q) . . . . .	16
4.7.1	Équation de Yule-Walker . . . . .	16
<b>5</b>	<b>Les processus ARMA(<math>p, q</math>)</b>	<b>17</b>
5.1	ARMA(p,q) . . . . .	17
5.2	ARMA(1,1) . . . . .	17
<b>6</b>	<b>Modèle Linéaire non Stationnaire</b>	<b>18</b>
6.1	Marche Aléatoire . . . . .	18
6.2	Tendance Linéaire . . . . .	18
<b>7</b>	<b>Modèle ARIMA(<math>p, d, q</math>)</b>	<b>19</b>
7.1	Méthode de Box-Jenin . . . . .	19
7.1.1	Identification . . . . .	19
7.1.2	Estimation des paramètres . . . . .	19
7.1.3	Verification . . . . .	19

<b>8 Application à l'indice du CAC40</b>	<b>20</b>
8.1 Identification avec $\Delta^1$	20
8.1.1 Données	20
8.1.2 Stationnarité	20
8.1.3 Paramètres p	21
8.1.4 Paramètres q	22
8.1.5 Autocorrélation partielle trop significative	22
8.2 Identification avec $\Delta^2$	23
8.2.1 Stationnarité	23
8.2.2 Estimation de p et q	24
8.3 Estimations des paramètres	24
8.4 Vérification	25
8.4.1 Graphique	25
8.4.2 Moyenne	25
8.5 Prédiction	26
8.6 Interprétation et Conclusion	26

# 1 Introduction aux Séries Temporelles

## 1.1 Généralités

Nous allons dans un premier temps introduire les séries temporelles. L'étude des séries temporelles correspond à l'analyse statistique d'observations régulièrement espacées dans le temps. On les note  $(X_t)_{1 \leq t \leq T}$  où  $x_t$  correspond à la  $t$ -ième observation de  $X$ .

Nous pouvons nous demander s'il est possible de créer un lien entre les différentes valeurs de  $X_t$ . Nous pouvons décomposer la série temporelle  $(X_t)$  comme :

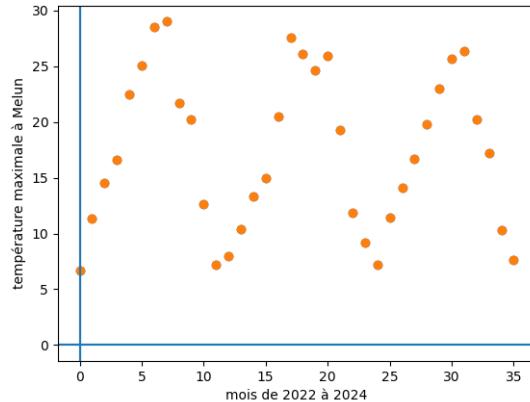
1. Une tendance  $T_t$  qui correspond à l'évolution long terme de la série. Il existe plusieurs types de tendance tels que les tendances linéaires, quadratiques, logarithmiques...
2. Une saisonnalité  $S_t$  qui correspond à un phénomène périodique
3. Une erreur  $\epsilon_t$  qui correspond à une partie aléatoire (stationnaire ou non)

La décomposition de la série peut-être additive :  $X_t = T_t + S_t + \epsilon_t$ ,  
multiplicative :  $X_t = T_t \times S_t \times \epsilon_t$ ,  
ou autre comme par exemple la décomposition suivante :  $X_t = T_t \times (S_t + \epsilon_t)$

**Hypothèses :** On suppose que les erreurs  $\epsilon_t$  sont :

- centrés i.e.  $\mathbb{E}[\epsilon_t] = 0$
- homoscédastiques i.e.  $V[\epsilon_t] = \sigma^2 \forall t$
- non auto-corrélatifs i.e.  $Cov[\epsilon_i, \epsilon_j] = 0 \forall i \neq j$

**Exemple :** En prenant les données de météofrance (1), nous pouvons construire une série temporelle  $(M_t)_{1 \leq t \leq 36}$  qui nous donne les températures maximales par mois observées à Melun de 2022 à 2024.



En reprenant l'exemple des séries temporelles climatiques qu'on note  $M_t$ , on suppose que la tendance  $T_t$  est positive due au dérèglement climatique ; cette tendance semble, par ailleurs, être linéaire à court terme (2). La saisonnalité  $S_t$  correspond au cycle d'une année et l'erreur  $\epsilon_t$  correspond aux variations des températures d'une année à l'autre dues à d'autres facteurs.

On suppose que  $T_t$  et  $S_t$  sont des combinaisons linéaires de fonction  $T_t^i$  et  $S_t^j$  c'est-à-dire :

$$\begin{cases} T_t = \alpha_1 T_t^1 + \alpha_2 T_t^2 + \dots + \alpha_n T_t^n \\ S_t = \beta_1 S_t^1 + \beta_2 S_t^2 + \dots + \beta_m S_t^m \end{cases}$$

Les fonctions  $T_t^i$  et  $S_t^j$  peuvent être des indicatrices ou d'autres types de fonction. Cela dépend principalement du type de données. Le but est de chercher à estimer les paramètres  $\alpha_i$  et  $\beta_j$  afin d'obtenir des estimations de la tendance  $T_t$  et de la saisonnalité  $S_t$ . Cela permettra ainsi d'obtenir une estimation de la série  $X_t$  (au temps  $t$ ).

**Notation :** Nous utiliserons les notations usuelles i.e. :

- $\hat{\alpha}_i$  (resp.  $\hat{\beta}_j$ ) est l'estimation du paramètre  $\alpha_i$  (resp.  $\beta_j$ )

- $\widehat{\alpha}$  (resp.  $\widehat{\beta}$ ) est l'estimation du vecteur  $\alpha$  (resp.  $\beta$ )
- $\widehat{\alpha} = (\widehat{\alpha}_1, \widehat{\alpha}_2, \dots, \widehat{\alpha}_n)^t$  et  $\widehat{\beta} = (\widehat{\beta}_1, \widehat{\beta}_2, \dots, \widehat{\beta}_m)^t$   $t$  correspond à la transposé et non au temps
- $\widehat{S}_t$  (resp.  $\widehat{T}_t$ ) correspond à l'estimation de  $S_t$  (resp.  $T_t$ )
- $\widehat{X}_t$  correspond à l'estimation de  $X_t$

Si la décomposition de la série est additive alors,  $X_t = \widehat{X}_t + \epsilon_t$ .

## 1.2 Estimateur des Moindres Carrés Ordinaires

Nous considérons le modèle où la décomposition de la série est additive. On a ainsi :

$$\begin{aligned} X_t = T_t + S_t + \epsilon_t &\Leftrightarrow X_t = \alpha_1 T_t^1 + \alpha_2 T_t^2 + \dots + \alpha_n T_t^n + \beta_1 S_t^1 + \beta_2 S_t^2 + \dots + \beta_m S_t^m + \epsilon_t \\ &\Leftrightarrow X_t = \sum_{i=1}^n \alpha_i T_t^i + \sum_{j=1}^m \beta_j S_t^j + \epsilon_t \\ &\Leftrightarrow \epsilon_t = X_t - \sum_{i=1}^n \alpha_i T_t^i - \sum_{j=1}^m \beta_j S_t^j \end{aligned}$$

On cherche à minimiser la somme des résidus  $\epsilon_t$  donc :

$$\arg \min \{ \sum_{t=1}^T \epsilon_t^2 \} = \arg \min \{ \sum_{t=1}^T (X_t - \sum_{i=1}^n \alpha_i T_t^i - \sum_{j=1}^m \beta_j S_t^j)^2 \}$$

$$\mathbf{T} = \begin{bmatrix} | & | & | \\ T^1 & T^2 & \dots & T^n \\ | & | & & | \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} | & | & | \\ S^1 & S^2 & \dots & S^m \\ | & | & & | \end{bmatrix}$$

On a donc le modèle suivant :  $\mathbf{X} = \mathbf{T}\alpha + \mathbf{S}\beta + \epsilon$ . En posant :

$$\mathbf{Y} = [\mathbf{T} \quad | \quad \mathbf{S}] \quad \nu = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

On obtient le modèle suivant :  $\mathbf{X} = \mathbf{Y}\nu + \epsilon$

En utilisant l'estimateur des moindre carrés ordinaire ;  $\mathbf{Y}'\mathbf{Y}\widehat{\nu} = \mathbf{Y}'\mathbf{X}$ .

$$\text{donc} : \widehat{\nu} = (\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{X} \Leftrightarrow \widehat{\nu} = \left[ \begin{bmatrix} \mathbf{T} \\ \mathbf{S} \end{bmatrix} \right] \left[ \begin{bmatrix} \mathbf{T} & \mathbf{S} \end{bmatrix} \right]^{-1} \left[ \begin{bmatrix} \mathbf{T} \\ \mathbf{S} \end{bmatrix} \right] \mathbf{X}$$

**Remarque :** Si il n'y a pas de saisonnalité, on retombe sur l'estimateur MCO classique avec :  $\widehat{\beta} = [\mathbf{T}'\mathbf{T}]\mathbf{T}'\mathbf{X}$ .

## 1.3 Tendance

**Définition :** La tendance d'une série temporelle correspond à une évolution à long terme de la série.

Comme expliqué précédemment, la tendance peut-être linéaire, quadratique où logarithmique. Il existe cependant d'autres type de tendance.

- tendance linéaire :  $T_t = bt + a$  où  $a, b \in \mathbb{R}$
- tendance quadratique :  $T_t = ct^2 + bt + a$  où  $a, b, c \in \mathbb{R}$
- tendance puissance n :  $T_t = a_0 + a_1 t + \dots + a_n t^n$  où  $a_0, \dots, a_n \in \mathbb{R}$
- tendance logarithmique :  $T_t = a \log(t)$  où  $a \in \mathbb{R}$
- tendance exponentiel :  $T_t = a(1+k)^t$  où  $a \in \mathbb{R}, k > 0$
- tendance de Gompertz :  $T_t = e^{a+kb^t}$  où  $a, b, k \in \mathbb{R}$

Cette tendance qu'on note  $T_t$  permet, comme son nom l'indique, de déterminer comment la série évolue à long terme. Comme expliquer précédemment, on suppose que la tendance  $T_t$  est une combinaison linéaire de fonctions  $T_t^i$ .

$$T_t = \sum_{i=1}^n \alpha_i T_t^i$$

**Exemple :** Si la tendance est quadratique alors  $T_t^1 = 1, T_t^2 = t, T_t^3 = t^2$  et ainsi :

$$\begin{aligned} T_t &= \sum_{i=1}^3 \alpha_i T_t^i \\ &= \alpha_1 T_t^1 + \alpha_2 T_t^2 + \alpha_3 T_t^3 \\ &= \alpha_1 + \alpha_2 t + \alpha_3 t^2 \end{aligned}$$

Cette tendance quadratique peut ainsi être représenté sur un graphique.

**Remarque :** Une tendance linéaire indique une croissance constante, une tendance exponentielle indique

une croissance de plus en plus importante, inverse pour la tendance logarithmique. Pour ce qui est de la tendance quadratique, il est plus compliqué de trouver le "sens" de celle-ci.

## 1.4 Saisonnalité

**Définition :** La saisonnalité d'une série temporelle correspond à un phénomène périodique identifié.

On rappelle que la saisonnalité  $S_t$  est une combinaison linéaire des fonctions  $S_t^j$ .

$$S_t = \sum_{j=1}^m \beta_j S_t^j$$

On considérera que les fonctions  $S_t^j$  sont des indicatrices.

$$S_t^j : \begin{cases} \mathbb{N} \rightarrow \{0, 1\} \\ t \rightarrow \begin{cases} 1 & \text{sous une condition temporelle} \\ 0 & \text{sinon} \end{cases} \end{cases}$$

### 1.4.1 Modèle trimestriel de Buys-Ballot

Ce modèle est un exemple d'un type de saisonnalité qui est trimestrielle.

Chaque fonction  $S_t^j$  est l'indicatrice du trimestre  $j$  avec  $j \in \{1, 2, 3, 4\}$

$$S_t = \beta_1 S_t^1 + \beta_2 S_t^2 + \beta_3 S_t^3 + \beta_4 S_t^4$$

Dans ce modèle, si les séries temporelles sont des données trimestrielles, on a :

$$S_t^1 = \begin{cases} 1 & \text{si } t = \text{trimestre 1} \\ 0 & \text{sinon} \end{cases} \quad \dots \quad S_t^4 = \begin{cases} 1 & \text{si } t = \text{trimestre 4} \\ 0 & \text{sinon} \end{cases}$$

**Remarque :** Dans ce modèle de Buys-Ballot, la tendance est supposé linéaire ce que donne l'expression suivante. On suppose la décomposition de la série additive et on obtient le modèle suivant :

$$X_t = \underbrace{\alpha_1 + \alpha_2 t}_{T_t} + \underbrace{\beta_1 S_t^1 + \beta_2 S_t^2 + \beta_3 S_t^3 + \beta_4 S_t^4}_{S_t} + \epsilon_t$$

Par exemple, lorsque l'on est au premier trimestre, on a :  $S_t^1 = 1$  et  $S_t^2 = S_t^3 = S_t^4 = 0$ .

L'expression devient alors :

$$X_t = \underbrace{\alpha_1 + \alpha_2 t}_{T_t} + \underbrace{\beta_1}_{S_t} + \epsilon_t$$

## 1.5 Bruit Blanc

Il existe deux manières de définir le bruit blanc. Le bruit blanc correspond à l'erreur.

**Bruit Blanc Faible :** On appelle bruit blanc faible toute suite  $(\epsilon_t)_{t \in \mathbb{Z}}$  tel que :

- $\mathbb{E}(\epsilon_t) = 0$
- $V(\epsilon_t) = \sigma^2$
- $Cov(\epsilon_{t_1}, \epsilon_{t_2}) = 0 \forall t_1 \neq t_2$

**Bruit Blanc Fort :** On appelle bruit blanc faible toute suite  $(\epsilon_t)_{t \in \mathbb{Z}}$  tel que  $(\epsilon_t)$  soit i.i.d (i.e. des variables indépendantes et identiquement distribuées.)

**Remarque :** Lorsque le bruit blanc est fort :

$$\epsilon_t \sim \mathcal{N}(0, \sigma^2)$$

## 2 Moyenne Mobile

### 2.1 Généralité

**Définition :** Une moyenne mobile permet d'extraire les composantes basses fréquentes de la série (i.e. la tendance).

Les moyennes mobiles permettent de lisser les séries. Elles conservent la tendance en éliminant le cycle.

On se place dans  $\mathbb{C}^{\mathbb{N}}$  un espace vectoriel et on introduit les opérateurs suivants :

$$L : X_t \rightarrow X_{t-1}$$

$$F : X_t \rightarrow X_{t+1}$$

**Remarque :**  $\mathbb{C}^{\mathbb{N}}$  correspond à l'espace vectoriel des suites à valeurs complexes.

L'opérateur  $L$  est appelé opérateur de retard et l'opérateur  $F$ , l'opérateur d'avance.  
Ces opérateurs sont linéaires.

#### Démonstration

- Soit  $(X_t)_{t \in \mathbb{N}}$  et  $(Y_t)_{t \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$ , deux séries temporelles.  
 $(X_t)_{t \in \mathbb{N}} + (Y_t)_{t \in \mathbb{N}} = (X_t + Y_t)_{t \in \mathbb{N}}$

$$\begin{aligned} L((X_t + Y_t)_{t \in \mathbb{N}}) &= (X_{t-1} + Y_{t-1})_{t \in \mathbb{N}} \\ &= (X_{t-1})_{t \in \mathbb{N}} + (Y_{t-1})_{t \in \mathbb{N}} \\ &= L((X_t)_{t \in \mathbb{N}}) + L((Y_t)_{t \in \mathbb{N}}) \end{aligned}$$

- Soit  $(X_t)_{t \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$ , une série temporelle et  $\lambda \in \mathbb{C}$   
 $\lambda \cdot (X_t)_{t \in \mathbb{N}} = (\lambda X_t)_{t \in \mathbb{N}}$

$$\begin{aligned} L((\lambda X_t)_{t \in \mathbb{N}}) &= (\lambda X_{t-1})_{t \in \mathbb{N}} \\ &= \lambda \cdot (X_{t-1})_{t \in \mathbb{N}} \\ &= \lambda \cdot L((X_t)_{t \in \mathbb{N}}) \end{aligned}$$

- $\forall (X_t)_{t \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}}, L((X_t)_{t \in \mathbb{N}}) \in \mathbb{C}^{\mathbb{N}}$

On en déduit ainsi que l'opérateur  $L$  est linéaire. On peut montrer la linéarité de l'opérateur  $F$  de la même manière.

**Remarque :**  $L$  et  $F$  sont des opérateurs inverses. C'est à dire que  $L^{-1} = F$  et  $F^{-1} = L$ .

On a ainsi la relation :  $F \circ L = L \circ F = \mathbb{I}_d$

#### Démonstration

$$F \circ L = F \circ F^{-1} = F^0 = \mathbb{I}_d = L^0 = L \circ L^{-1} = L \circ F$$

**Rappel :** Un opérateur à une puissance  $n$  est la composé de l'opérateur  $n$  fois.

$$L^n = \underbrace{L \circ L \circ \dots \circ L}_{n \text{ fois}}$$

On peut ainsi créer des polynomes annulateur de ces opérateurs. Sans rentré dans les détails car cela rentre dans le cadre de mon projet tuteuré de L2, on a :

$$P(X) = a_0 X^0 + a_1 X^1 + a_2 X^2 + \dots + a_n X^n$$

$$P(L) = a_0 \mathbb{I}_d + a_1 L + a_2 L^2 + \dots + a_n L^n$$

**Propriétés :** Soit  $A$  et  $B$  deux polynômes qui sont des moyennes mobiles

- $A(L) + B(L) = (A + B)(L)$
- $\forall \alpha \in \mathbb{C}, \alpha \cdot A(L) = (\alpha A)(L)$
- $A(L) \circ B(L) = (AB)(L) = B(L) \circ A(L)$

**Remarque :** Ces polynomes permettent de diagonaliser et de trouver les valeurs propres grâce aux théorème de Cayley-Hamilton. Nous n'utiliserons pas dans ce projet les valeurs propres.

## 2.2 Utilisation des Moyennes Mobiles

**Définition :** Une moyenne mobile est un opérateur qui est une combinaison linéaire de l'opérateur retard.

**Exemple :** La moyenne mobile la plus simple est la moyenne mobile centrée et symétrique. Cette moyenne est centrée sur la valeur qu'on notera  $t_0$  et prend les  $m$  valeurs précédentes et les  $m$  valeurs suivantes. On la note  $M_m$

$$M_m = \frac{1}{2m+1} (X_{t_0-m} + X_{t_0-m+1} + \dots + X_{t_0-1} + X_{t_0} + X_{t_0+1} + \dots + X_{t_0+m-1} + X_{t_0+m})$$

Soit la série  $(X_t)_{t \in \mathbb{N}^*}$  définie par  $X_t = 1, 2 + 0.2t + 0.1S_t^1 - 0.05S_t^2 + 1.5S_t^3 - 0.2S_t^4$ .

Dans cet exemple,  $t$  correspond au  $t$ -ième mois, la première valeur de  $t$  est le mois de janvier. Les trimestres correspondent aux mois de janvier-février-mars, ..., octobre-novembre-décembre

Les fonctions  $S_t^j$  sont les fonctions indicatrices du trimestres  $j$ .

On fixe  $t_0 = 10$  et  $m = 2$

On peut ainsi calculer la moyenne mobile en ce point.

$X_8 = 6, 7, X_9 = 7, 2, X_{10} = 6, X_{11} = 6, 5, X_{12} = 7$

$$\begin{aligned} M_2 &= \frac{1}{2 \times 2 + 1} (X_{10-2} + X_{10-1} + X_{10} + X_{10+1} + X_{10+2}) \\ &= \frac{1}{5} (6, 7 + 7, 2 + 6 + 6, 5 + 7) \\ &= 6, 68 \end{aligned}$$

**Remarque :** Cette moyenne mobile correspond à un exemple d'une moyenne mobile symétrique et centrée. Il existe d'autres types de moyenne mobile qui ne sont pas nécessairement centrés ou symétriques.

De manière plus générale, une moyenne mobile s'écrit :

$$M = \sum_{i=-m_1}^{m_2} \lambda_i L^i$$

avec  $m_1, m_2 \in \mathbb{N}$

**Définition :** Si  $m_1 = m_2 = m$ , on dit que la moyenne  $M$  est centrée.

**Définition :** Si  $M$  est centré et  $\forall i \{1, \dots, m\}, \lambda_i = \lambda_{-i}$ , on dit que  $M$  est symétrique.

**Remarque :** Les moyennes mobiles centrées symétriques sont d'ordre impair (nombre de termes dans la moyenne). Il est possible de construire artificiellement une série d'ordre pair.

## 2.3 Valeurs Propres

**Définition :** Une série  $(X_t)_{t \in \mathbb{N}}$  est dite absorbée par  $M$  si  $\forall t, M(X_t) = 0$

**Exemple :** La série  $X_t = -X_{t-1} - X_{t-2}$  est absorbée par la moyenne mobile  $M(X_t) = X_{t+1} + X_t + X_{t-1}$

**Proposition :** Les vecteurs propres associés à  $\lambda = 0$  forment un espace vectoriel de dimension  $m_1 + m_2$ . Une base est constituée des vecteurs  $\lambda_i r^t$

Dans notre exemple, on peut trouver une base de vecteurs propre en trouvant les racines de :

$$r^{t+1} + r^t + r^{t-1} = 0 \iff r^2 + r + 1 = 0$$

$$\implies r = \frac{-1 \pm i\sqrt{3}}{2}$$

Les solutions de l'équation sont ainsi :

$$\begin{aligned} r_1^t &= \cos\left(\frac{2t\pi}{3}\right) + i\sin\left(\frac{2t\pi}{3}\right) \\ r_2^t &= \cos\left(\frac{2t\pi}{3}\right) - i\sin\left(\frac{2t\pi}{3}\right) \end{aligned}$$

Les séries absorbées par la moyenne sont donc de la formes :

$$X_t = \alpha r_1^t + \beta r_2^t$$

L'ensemble de ces séries peut-être engendrés par la famille de vecteur :

$$B = \{\cos\left(\frac{2t\pi}{3}\right), \sin\left(\frac{2t\pi}{3}\right)\}$$

La famille  $B$  est également libre. Il s'agit donc d'une base de vecteurs propres. Les séries absorbées par  $M$  sont donc :

$$X_t = \lambda \cos\left(\frac{2t\pi}{3}\right) + \nu \sin\left(\frac{2t\pi}{3}\right)$$

## 2.4 Exemple de Moyennes Mobiles

### 2.4.1 Forme générale

La forme générale d'une moyenne mobile est de la forme :

$$M = \sum_{i=-m_1}^{m_2} \lambda_i L^i$$

$m_1$  et  $m_2$  correspondent aux "bornes" de notre moyenne mobile.  $L^i$  correpond à l'opérateur retard donc à une valeurs de  $X_t$  qu'on multiplie par une scalaire  $\lambda_i$  de notre corps  $\mathbb{K}$  (usuellement  $\mathbb{R}$  ou  $\mathbb{C}$ ).

### 2.4.2 Moyenne Mobile Paire

Cette moyenne mobile est centrée. On la définit comme :

$$M = \sum_{i=-m}^m \lambda_i L^i \quad \begin{cases} \lambda_i = \frac{1}{2m} & \text{si } |i| \neq m \\ \lambda_i = \frac{1}{4m} & \text{si } |i| = m \end{cases}$$

**Exemple :** Lorsque  $m = 2$ , la moyenne mobile  $M$  vaut :

$$M(X_t) = \frac{1}{8}X_{t-2} + \frac{1}{4}X_{t-1} + \frac{1}{4}X_t + \frac{1}{4}X_{t+1} + \frac{1}{8}X_{t+2}$$

### 2.4.3 Moyenne Mobile Impaire

Les moyennes mobiles impaires sont de la même forme que les moyennes paires mises à part au niveau des scalaires. On a ainsi :

$$M = \sum_{i=-m}^m \lambda_i L^i \quad \lambda_i = \frac{1}{2m+1} \iff M = \frac{1}{2m+1} \sum_{i=-m}^m L^i$$

### 2.4.4 Moyenne Mobile Différenciée

La moyenne mobile différenciée se note  $\Delta_p$  elle vaut :

$$\Delta_p = (\mathbb{I}_d - L)^p$$

L'expression de cette moyenne est ainsi :

$$\Delta_p = \sum_{i=0}^p \binom{i}{p} L^i (X_t) \iff \Delta_p = \sum_{i=0}^p \binom{i}{p} X_{t-i}$$

### 3 Lissage Exponentiel

Le but d'un lissage est d'estimer l'évolution de la série en un temps  $T + 1$ . Il est possible d'effectuer des lissages court, moyen ou long terme.

#### 3.1 Lissage Exponentiel Simple

On dispose d'une série  $(X_t)_{1 \leq t \leq N}$ , correspondant aux observations aux temps  $1, \dots, N$ . La prévision de  $X_t$  à un horizon  $h$  est fourni par le modèle suivant :

$$\widehat{X}_t(h) = (1 - \beta) \sum_{i=0}^{t-1} \beta^i X_{t-i} \quad \beta \in ]0, 1[$$

$\beta$  est une constante de lissage. Un  $\beta$  proche de 1 indique une prise en compte importante du passé tandis qu'un  $\beta$  proche de 0 indique une prise en compte faible du passé.

**Remarque :** On note  $F_{T+1}$  la valeur prédictive en  $T + 1$

**Proposition :** On a :  $1 - \beta = \alpha$

##### 3.1.1 Ordre 1

$$\widehat{X}_T = (1 - \beta)X_T + \beta\widehat{X}_{T-1}$$

Comme on a :  $1 - \beta = \alpha$  et  $F_N = \widehat{X}_{N-1}$ , on peut écrire la relation à l'ordre 1 de la manière suivante :

$$F_{T+1} = \alpha X_T + (1 - \alpha)F_T$$

##### 3.1.2 Itération

$$\begin{cases} F_0 = X_1 \text{ ou } (X_1 + X_2 + \dots + X_p)/p \\ F_{t+1} = \alpha X_t + (1 - \alpha)F_t \text{ si } t \leq N \\ F_{t+1} = F_{N+1} \text{ sinon} \end{cases}$$

### 3.2 Lissage Exponentiel Double

Avec un lissage exponentiel double, on approche la série par une droite. La prévision à l'horizon  $h$  est la suivante :

$$F_{T+h} = \widehat{X}_T(h) = \widehat{A}(T) + h\widehat{B}(T)$$

On a :

$$\begin{aligned} \widehat{A}(T) &= 2S_1(T) - S_2(T) \\ \widehat{B}(T) &= \frac{1-\beta}{\beta}(S_1(T) - S_2(T)) \end{aligned}$$

Pour trouver  $A$  et  $B$ , il faut minimiser l'expression suivante :

$$\min_{A,B} \left\{ \sum_{i=0}^{T-1} \beta^i (X_{T-i} - [A + (T-i)B])^2 \right\}$$

D'après la démonstration de Gouriéroux et Montfort :

$$\begin{cases} S_1(t) = (1 - \beta)X_t + \beta S_1(t-1) \text{ série lissée une fois} \\ S_2(t) = (1 - \beta)^2 \sum_{k=0}^{t-1} \sum_{i=0}^{t-k-1} \beta^{i+k} X_{t-(k+i)} \text{ série lissée deux fois} \end{cases}$$

##### 3.2.1 Ordre 1

Si en  $T$ , on a :  $F_{T+1} = \widehat{X}_T(1) = \widehat{A}(T) + \widehat{B}(T)$  alors,

$$\begin{cases} A(T+1) = (1 - \beta^2)[X_{T+1} + \widehat{X}_T(1)] + \widehat{A}(T) + \widehat{B}(T) \\ B(T+1) = \widehat{B}(T) + (1 - \beta^2)[X_{T+1} + \widehat{X}_T(1)] \end{cases}$$

### 3.2.2 Formule classique

La formule classique permet d'obtenir la série lissée à un horizon  $h$ , elle est donnée par la formule suivante où  $\alpha \in ]0, 1[$

$$A_j = 2\hat{y}_j^1 - \hat{y}_j^2$$

$$B_j = \frac{\alpha}{1-\alpha}[\hat{y}_j^1 - \hat{y}_j^2]$$

**Remarque :** On note  $\hat{y}_j^1$  la série lissée une fois et  $\hat{y}_j^2$  la série lissée deux fois (il ne s'agit pas de puissance mais d'indice)

$$\begin{cases} \hat{y}_j^1 = \alpha y_j + (1-\alpha)\hat{y}_{j-1}^1 \\ \hat{y}_j^2 = \alpha\hat{y}_j^1 + (1-\alpha)\hat{y}_{j-1}^2 \end{cases}$$

### 3.2.3 Itération

On peut construire le lissage exponentiel double par itération de la manière suivante :

$$\begin{cases} S_0^1 = X_1 \text{ ou } (X_1 + \dots + X_p)/p \\ S_{t+1}^1 = \alpha X_t + (1-\alpha)S_t^1 \text{ si } 0 \leq t \leq N \\ S_0^2 = 0 \\ S_{t+1}^2 = \alpha S_t^1 + (1-\alpha)S_t^2 \text{ si } 0 \leq t \leq N \\ A_{t+1} = 2S_{t+1}^1 - S_{t+1}^2 \text{ si } 0 \leq t \leq N \\ B_{t+1} = \frac{\alpha}{1-\alpha}[S_{t+1}^1 - S_{t+1}^2] \text{ si } 0 \leq t \leq N \\ F_{t+1} = A_{t+1} + B_{t+1} \text{ si } 0 \leq t \leq N \\ F_t = A_{N+1} + (t-N-1)B_{N+1} \end{cases}$$

## 4 Introduction au Modèle ARIMA

### 4.1 Les Espaces $\mathcal{L}^2$

**Définition :** L'espace  $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$  est l'espace des variables de carré intégrable.

$\mathcal{L}^n$  est l'espace de Banach des classes équivalentes pour l'égalité  $\mathbb{P}$ -presque sûre des fonctions mesurables i.e.  $\|f\|_n = |\int |f|^n d\mathbb{P}|$  finie.

On considère le processus  $(X_t)$  défini sur l'espace de probabilité  $\mathcal{L}^2$  à valeur dans  $\mathbb{R}$ .

**Remarque :** Dans l'espace  $\mathcal{L}^2$ , les variances/covariances sont finies.

**Proposition :**  $\mathcal{L}^2$  est un espace de Hilbert munie du produit scalaire  $\langle \cdot, \cdot \rangle$  et de la norme associé  $\|\cdot\|$

$$\begin{cases} \langle X, Y \rangle = \mathbb{E}[XY] \\ \|X\| = \sqrt{\mathbb{E}[X^2]} \end{cases}$$

**Théorème de Projection :** Si  $F$  est un fermé de  $\mathcal{L}^2$  (i.e.  $\mathcal{L}^2 \setminus F$  ouvert) alors,  $\forall Y \in \mathcal{L}^2, \exists! \hat{Y} \in F$  tel que :

$$\|Y - \hat{Y}\| = \min_{f \in F} \{ \|Y - f\| \}$$

**Remarque :**  $\hat{Y} \in F \subset \mathcal{L}^2$  donc  $\hat{Y}$  est une variable aléatoire.

**Définition :**  $X_n$  converge vers  $X$  dans  $\mathcal{L}^2$  si :

$$\lim_{n \rightarrow +\infty} \|X_n - X\| = 0 \text{ c'est à dire } \begin{cases} \mathbb{E}[X_n] \xrightarrow{n \rightarrow +\infty} \mathbb{E}[X] \\ \mathbb{V}[X_n - X] \xrightarrow{n \rightarrow +\infty} 0 \end{cases}$$

**Définition :** On définit ainsi la variable aléatoire  $Y$ , comme limite dans  $\mathcal{L}^2$  de  $Y_{p,q}$

$$Y_{p,q} = \sum_{n=-p}^q a_n X_n$$

$$Y = \lim_{p,q \rightarrow +\infty} Y_{p,q} \implies Y = \sum_{n \in \mathbb{N}} a_n X_n$$

### 4.2 Vecteurs et Processus Gaussiens

**Rappels :** Pour un vecteur  $A \in \mathbb{R}^n$  où  $A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ , on a :

$$\mathbb{E}[A] = \begin{pmatrix} \mathbb{E}[a_1] \\ \mathbb{E}[a_2] \\ \vdots \\ \mathbb{E}[a_n] \end{pmatrix}$$

$$\mathbb{V}[A] = \mathbb{E}[(A - \mathbb{E}[A])(A - \mathbb{E}[A])']$$

$\mathbb{V}[A]$  est la matrice de variance-covariance (on admet qu'elle existe). Cette matrice est hermitienne c'est à dire égale à sa transposée conjuguée.

**Proposition :** Soit  $f$  une application linéaire et  $M$  sa matrice associée. On a alors :

$$\begin{cases} \mathbb{E}[MA] = M\mathbb{E}[A] \\ \mathbb{V}[MA] = M\mathbb{V}[A]M' \end{cases}$$

### 4.2.1 Vecteurs Gaussiens

**Définition :** Le vecteur  $A = (A_1, \dots, A_n)'$  est un vecteur gaussien si toute combinaison des  $A_i$  est une variable gaussienne i.e.  $\forall \alpha \in \mathbb{R}^n$ ,  $\alpha A$  est une variable gaussienne.

On note sa densité :

$$f(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{\det \Sigma}} \exp \left( -\frac{1}{2}(x - \mu)' \Sigma^{-1} (x - \mu) \right)$$

On a  $\mu \in \mathbb{R}^n$ ,  $\Sigma$  une matrice hermitienne positive  $n \times n$

**Remarque :** Si  $A$  est un vecteur gaussien alors  $\mathbb{E}[A] = \mu$  et  $\mathbb{V}[A] = \Sigma$

### 4.2.2 Processus Gaussien

**Définition :** On dit que  $(X_t)_{1 \leq t \leq T}$  est un processus gaussien si tout système fini extrait est un vecteur gaussien i.e.

$$\forall p < T \in \mathbb{N}, \forall (t_1, \dots, t_p) \in \mathbb{R}^p, (X_{t_1}, \dots, X_{t_p}) \text{ est un vecteur gaussien}$$

## 4.3 Régression dans $\mathcal{L}^2$

### 4.3.1 Nombre fini de variables

On cherche à faire une régression affine de  $Y$  sur  $X_1, \dots, X_n$ , la projection orthogonale dans  $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$  de  $Y$  sur  $\mathcal{H} = \text{Vect}(\mathbb{I}_d, X_1, \dots, X_n)$ .

$$\hat{Y} = EL(Y | \mathbb{I}_d, X_1, \dots, X_n)$$

$EL$  est l'espérance linéaire de  $Y$  sachant les variables  $\mathbb{I}_d, X_1, \dots, X_n$

**Proposition :**  $EL(Y | \mathbb{I}_d, X_1, \dots, X_n)$  est la meilleure approximation de  $Y$  par une combinaison linéaire de  $\mathbb{I}_d, X_1, \dots, X_n$  dans  $\mathcal{L}^2$

**Proposition :** Soit  $\eta = [\text{Cov}(X, X_i)]_{i=0,1,\dots,n}$  et  $\Sigma = [\text{Cov}(X_i, X_j)]_{i,j=0,1,\dots,n}$ .

$$\hat{X} = EL(Y | \mathbb{I}_d, X_1, \dots, X_n) = a_0 + a_1 X_1 + \dots + a_n X_n$$

Le vecteur  $A = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}$  vérifie  $A = \Sigma^{-1} \eta$

**Remarque :**  $X - \hat{X} \in \mathcal{H}^\perp$  donc  $\mathbb{E}[\hat{X}] = \mathbb{E}[X]$

### 4.3.2 Nombre infini de variables

On considère  $X_1, \dots, X_n, \dots$  une infinité de variables. Soit  $\mathcal{H}$  l'adhérence des combinaisons linéaires des variables.

Pour rappel l'adhérence d'un ensemble  $Z$  est :  $\overline{Z} = \bigcup_{\substack{O \subset Z \\ O \text{ ouvert}}} O$

On considère  $\widehat{X}_n = EL(Y | \mathbb{I}_d, X_1, \dots, X_n)$

$$\widehat{X} = \lim_{n \rightarrow +\infty} \widehat{X}_n = \lim_{n \rightarrow +\infty} EL(Y | \mathbb{I}_d, X_1, \dots, X_n)$$

### 4.3.3 Opérateur de Projection Linéaire

Soit  $Y$  une variable aléatoire et  $\{X_1, \dots, X_n\}$  une famille de  $n$  variables aléatoires, l'opérateur  $\Pi$  lui associe :

$$\Pi[Y | \{X_1, \dots, X_n\}] = \alpha_1 X_1 + \dots + \alpha_n X_n$$

Les  $\alpha_i$  sont obtenus en minimisant :

$$(\alpha_1, \dots, \alpha_n) = \arg \min_{\alpha_1, \dots, \alpha_n} \{\mathbb{V}[Y - (\alpha_1 X_1 + \dots + \alpha_n X_n)]\}$$

**Remarque :** L'opérateur  $\Pi$  est linéaire et l'erreur  $Y - \Pi[Y|\{X_1, \dots, X_n\}]$  est non corrélée au  $(X_i)$

**Proposition :**  $\text{Cov}(X_1, X_2) = 0 \implies \Pi[Y|\{X_1, X_2\}] = \Pi[Y|\{X_1\}] + \Pi[Y|\{X_2\}]$

#### 4.3.4 Prévision Linéaire

On définit le processus d'innovation du processus  $(X_t)$  est le processus  $(\epsilon_t)$  est définie par :

$$\epsilon_t = X_t - EL(X_t|\mathbb{I}_d, X_{t-1}, X_{t-2}, \dots)$$

avec :

$$EL(X_t|\mathbb{I}_d, X_{t-1}, X_{t-2}, \dots) = - \sum_{i=1}^{\infty} \rho^i X_{t-i} \text{ où } |\rho| < 1$$

**Propriété :** Soit  $(Y_t)$  un  $BB(0, \sigma^2)$ , le processus stationnaire  $(X_t)$  définit par :

$$X_t = Y_t - \rho Y_{t-1} \iff Y_t = X_t - \rho X_{t-1}$$

**Propriété :**  $\forall t \in \mathbb{N}, \forall k \geq 0, \forall h > 0 :$

$$\begin{cases} \mathbb{E}[\epsilon_t X_{t-h}] = 0 \\ \mathbb{E}[\epsilon_t X_t] \neq 0 \\ \mathbb{E}[\epsilon_{t+h} X_{t-h}] = 0 \end{cases}$$

#### 4.4 Polynômes Opérateur Retard et Avance

**Définition :** On définit  $P(L) = 1 - \lambda L$ , l'inverse de ce polynôme vaut alors :

$$(1 - \lambda L)^{-1} = \begin{cases} \sum_{k=0}^{+\infty} \lambda^k L^k \text{ si } |\lambda| < 1 \\ \sum_{k=1}^{+\infty} \frac{1}{\lambda^k} F^k \text{ si } |\lambda| > 1 \\ \text{non défini si } |\lambda| = 1 \end{cases}$$

Tout polynôme  $P(L)$  (normalisé tel que  $P(0) = 1$ ) peut sécrire :

$$P(L) = \prod_{i=1}^n (1 - \lambda_i L^i) \text{ avec } \lambda_i = \frac{1}{z_i}$$

$z_i$  est une racine du polynôme  $P$

**Proposition :**  $\forall i \in \{1, \dots, n\} |\lambda_i| \neq 1 \implies P(L)$  est inversible

#### 4.5 Les Séries Stationnaires

**Définition :** Un processus  $(X_t)_{1 \leq t \leq T}$  est stationnaire si :

$$\forall t \leq T, \mathbb{E}[X_t^2] < +\infty$$

$$\forall t \leq T, \mathbb{E}[X_t] = \mu$$

$$\forall t \leq T, \forall h \geq 0, \text{Cov}(X_{t+h}, X_t) = \gamma_X(h)$$

avec  $\gamma_X(h)$  une fonction indépendante de  $t$

**Définition :** On définit une fonction d'autocovariance et une fonction d'autocorrélation :

$$\gamma_X : h \longrightarrow \text{Cov}(X_{t+h}, X_t) \text{ fonction d'autocovariance}$$

$$\rho_X(h) : h \longrightarrow \frac{\text{Cov}(X_{t+h}, X_t)}{\sqrt{V(X_t)}\sqrt{V(X_{t-h})}} \text{ fonction d'autocorrélation}$$

**Définition :** On définit ainsi la matrice d'autocorrélation du vecteur  $(X_t, X_{t-1}, X_{t-2}, \dots, X_{t-h+1}, X_{t-h})$ , la matrice  $\mathcal{R}(h+1)$

$$\mathcal{R}(h+1) = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \rho(3) & \cdots & \rho(h) \\ \rho(1) & 1 & \rho(1) & \rho(2) & \cdots & \rho(h-1) \\ \rho(2) & \rho(1) & 1 & \rho(1) & \cdots & \rho(h-2) \\ \rho(3) & \rho(2) & \rho(1) & 1 & \cdots & \rho(h-3) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho(h) & \rho(h-1) & \rho(h-2) & \rho(h-3) & \cdots & 1 \end{pmatrix}$$

On a :

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}$$

L'estimation de  $\rho_X(h)$  est donnée par :

$$\widehat{\gamma}_X(h) = \frac{1}{T-h} \sum_{t=1}^{T-h} X_t X_{t+h}$$

On en déduit ainsi que :

$$\widehat{\rho}_X(h) = \frac{\widehat{\gamma}_X(h)}{\widehat{\gamma}_X(0)}$$

**Définition :** On définit pour une série stationnaire  $(X_t)$  la fonction d'autocorrélation partielle par :

$$\psi_X(h) : h \longrightarrow \widehat{\text{Corr}}(X_{t+h}, \widehat{X}_t)$$

où :

$$\begin{cases} \widehat{X}_{t-h} = X_{t-h} - EL(X_{t-h}|X_{t-1}, X_{t-2}, \dots, X_{t-h+1}) \\ \widehat{X}_t = X_t - EL(X_t|X_{t-1}, X_{t-2}, \dots, X_{t-h+1}) \end{cases}$$

## 4.6 Modèle AR(p)

Un modèle AR consiste à modéliser notre série temporelle à l'aide des termes passés. Il s'agit d'une régression de la variable par rapport à elle-même.

**Définition :** On appelle processus autorégressif d'ordre  $p$  noté  $AR(p)$ , un processus stationnaire suivant une relation :

$$\epsilon_t = X_t - \sum_{i=1}^p \phi_i X_{t-i}, \quad \forall t \in \mathbb{Z}$$

avec  $\phi_i \in \mathbb{R}$ ,  $(\epsilon_t)$  un BB de variance  $\sigma^2$

Cette expression est équivalente à l'expression suivante :

$$\epsilon_t = \Phi(L)X_t \quad \text{où } \Phi(L) = \mathbb{I}_d - \phi_1 L - \dots - \phi_p L^p$$

### 4.6.1 AR(1)

La forme général des processus de type AR(1) est :

$$X_t - \phi X_{t-1} = \epsilon_t \quad \text{où } \forall t \in \mathbb{Z}$$

$\epsilon_t$  est un BB de variance  $\sigma^2$

## 4.7 Modèle MA(q)

Le modèle MA consiste à modéliser notre série temporelle à l'aide des erreurs. Il s'agit d'une régression sur les erreurs.

Définition On écrit sous la forme MA( $\infty$ ) lorsque les racines de  $\Phi$  sont de module strictement inférieur à 1. On écrit alors sous la forme :

$$X_t = m + \sum_{i=0}^{+\infty} a_i \epsilon_{t-i}$$

avec  $a_0 = 1$ ,  $a_k \in \mathbb{R}$  et  $\sum_{i=0}^{+\infty} |a_i| < +\infty$

Lorsque certaines racines sont de module strictement inférieur à 1, on peut utiliser l'expression suivante :

$$\Phi * (z) = \left[ \prod_{j, |\lambda_j| < 1} (1 - \lambda_j z) \right] \left[ \prod_{j, |\lambda_j| > 1} \left( 1 - \frac{z}{\lambda_j} \right) \right]$$

**Remarque :** Soit  $(\eta_t)$ , le processus  $\eta_t = \Phi * (L) X_t$  est un bruit blanc. On appelle processus moyenne mobile d'ordre  $q$  noté MA(q) un processus stationnaire  $(X_t)$  vérifiant une relation :

$$X_t = \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i} \quad \forall t \in \mathbb{Z}$$

La fonction d'autocovariance est donnée par :

$$\gamma(h) = \mathbb{E}[X_t, X_{t-h}] = \begin{cases} [\theta_h + \theta_{h+1}\theta_1 + \dots + \theta_{h+q}\theta_q]\sigma^2 & \text{si } 1 \leq h \leq q \\ 0 & \text{si } h > q \end{cases}$$

Lorsque  $q = 1$ , le processus s'écrit :

$$X_t = \epsilon_t + \theta \epsilon_{t-1}$$

### 4.7.1 Équation de Yule-Walker

Le processus  $(X_t)$  s'écrit :

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \epsilon_t$$

$(X_t)$  est un processus AR(p) d'autocorrélation  $\rho(h)$ . On peut ainsi écrire le système sous forme matricielle :

$$\begin{pmatrix} \rho(1) \\ \rho(2) \\ \rho(3) \\ \vdots \\ \rho(p-1) \\ \rho(p) \end{pmatrix} = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(p-2) \\ \rho(2) & \rho(1) & 1 & \cdots & \rho(p-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho(p-2) & \rho(p-3) & \rho(p-4) & \cdots & \rho(1) \\ \rho(p-1) & \rho(p-2) & \rho(p-3) & \cdots & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_{p-1} \\ \phi_p \end{pmatrix}$$

**Remarque :** Le système d'équations engendré par cette équation matricielle est un système de Yule-Walker

## 5 Les processus $ARMA(p, q)$

### 5.1 ARMA(p,q)

Les processus ARMA généralisent les processus AR et MA. On appelle processus ARMA(p,q) un processus stationnaire ( $X_t$ ) vérifiant :

$$X_t - \sum_{i=1}^p \phi_i X_{t-i} = \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j} \quad \forall t \in \mathbb{Z}$$

Les  $\theta_i$  sont des réels et  $(\epsilon_t)$  un BB de variance  $\sigma^2$ . On peut aussi écrire le processus ARMA(p,q) sous la forme suivante :

$$\Phi(L)X_t = \Theta(L)\epsilon_t \text{ où } \begin{cases} \Phi(L) = \mathbb{I}_d - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \\ \Theta(L) = \mathbb{I}_d + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q \end{cases}$$

Le processus ARMA(p,q) peut ainsi s'écrire sous la forme :

- AR( $\infty$ ) en écrivant  $\epsilon_t = \Theta^{-1}(L)\Phi(L)X_t$
- MA( $\infty$ ) en écrivant  $X_t = \Phi^{-1}(L)\Theta(L)\epsilon_t$

**Propriété :** Les autocovariances  $\gamma(h)$  satisfont :

$$\gamma(h) - \sum_{i=1}^p \phi_i \gamma(h-i) = 0$$

**Propriété :** Les autocorrélations  $\gamma(h)$  satisfont :

$$\gamma(h) - \sum_{i=1}^p \phi_i \gamma(h-i) = \sigma^2[\theta_h + h_1 \theta_{h+1} + \dots + h_{q-h} \theta_q]$$

Les  $h_i$  correspondent aux coefficients de la forme MA( $\infty$ )

### 5.2 ARMA(1,1)

Soit  $(X_t)$  un processus ARMA(1,1), on a alors l'expression suivante :

$$X_t - \phi X_{t-1} = \epsilon_t + \theta \epsilon_{t-1}$$

avec  $\phi, \theta \in ]-1, 1[ \setminus \{0\}$

## 6 Modèle Linéaire non Stationnaire

Les résultats obtenus précédemment découlent de l'hypothèse de stationnarité. Dans cette partie nous nous intéresserons aux modèles qui ne respectent pas cette hypothèse.

### 6.1 Marche Aléatoire

On définit une marche aléatoire comme une suite de variable  $\epsilon_1, \dots, \epsilon_t$  une suite de variable *i.i.d.* de variance  $\sigma^2$  et d'espérance nulle. Une marche aléatoire vérifie :

$$Y_t = Y_{t-1} + \epsilon_t$$

Par convention,  $Y_0 = 0$

$$\begin{aligned}\mathbb{E}[Y_t] &= \mathbb{E}[Y_{t-1} + \epsilon_t] \\ &= \mathbb{E}[Y_{t-1}] + \mathbb{E}[\epsilon_t] \\ &= \mathbb{E}[Y_{t-2} + \epsilon_{t-1}] + 0 \\ &\quad \vdots \\ &= \mathbb{E}[Y_0] \\ &= 0\end{aligned}$$

$$\begin{aligned}V[Y_t] &= V[Y_{t-1} + \epsilon_t] \\ &= V[Y_0 + \sum_{i=1}^t e_i] \\ &= V[\sum_{i=1}^t e_i] \\ &= \sum_{i=1}^t V[e_i] \\ &= t\sigma^2\end{aligned}$$

### 6.2 Tendance Linéaire

Un processus non stationnaire linéaire est défini de la manière suivante,  $\epsilon_1, \dots, \epsilon_t$  une suite de variable *i.i.d.* de variance  $\sigma^2$  et d'espérance nulle qui vérifie :

$$Y_t = \alpha t + \epsilon_t$$

Ce processus a pour espérance et variance :

$$\begin{cases} \mathbb{E}[Y_t] = \alpha t \\ V[Y_t] = \sigma^2 \end{cases}$$

## 7 Modèle ARIMA(p,d,q)

Dans le modèle ARIMA(p,d,q), nous utiliserons l'opérateur moyenne mobile différence introduit précédemment. Pour rappel :

$$\Delta^d = (1 - L)^d$$

Un processus  $(X_t)$  est un processus ARIMA(p,d,q) autorégressif moyenne mobile intégré vérifiant :

$$\Phi(L)(1 - L)^d X_t = \Theta(L)\epsilon_t$$

où :

$$\begin{cases} \Phi(L) = \mathbb{I}_d - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \\ \Theta(L) = \mathbb{I}_d + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q \end{cases}$$

**Propriété :** Si  $(X_t)$  est un processus ARIMA(p,d,q) alors  $(\Delta^d X_t)$  converge vers un processus ARMA(p,q).

### 7.1 Méthode de Box-Jenin

L'estimation des paramètres de Box-Jenin se fait en 3 étapes. On choisit  $(X_t)$  une série temporelle donnant le prix d'un actif en fonction de  $t$  la temporalité.

#### 7.1.1 Identification

Dans un premier temps on observe si la série est stationnaire ou non. Pour une action c'est rarement le cas, celle-ci aura souvent une tendance haussière ou baissière. On devra donc la rendre stationnaire par différentiation, on devra donc déterminer la  $d$  du modèle d'ARIMA(p,d,q).

Une fois  $d$  déterminé, on devra déterminer les composantes  $p$  et  $q$  en utilisant les modèles  $AR(p)$  et  $MA(q)$ .

**Conclusion :** Grâce à l'identification, les composantes  $p, d, q$  sont déterminées.

#### 7.1.2 Estimation des paramètres

On rappelle que le modèle  $ARIMA(p, d, q)$  s'écrit sous la forme :

$$\Phi(L)(1 - L)^d X_t = \Theta(L)\epsilon_t$$

Nous connaissons maintenant  $p, d$  et  $q$ , nous cherchons maintenant à estimer les paramètre  $\phi_i$  et  $\theta_j$ , paramètres des modèles  $AR$  et  $MA$ .

#### 7.1.3 Verification

Une fois les paramètres déterminés, nous devons vérifier si celui-ci fonctionne ou non. Pour cela, nous avons choses à faire :

Premièrement, analyser les résidus pour voir s'ils ressemblent à du bruit blanc. Si ce n'est pas le cas, il faut modifier p,d et q.

Deuxièmement, il faut effectuer des tests statistiques de Ljung-Box afin de voir si les résidus présentent des autocorrélations.

## 8 Application à l'indice du CAC40

### 8.1 Identification avec $\Delta^1$

#### 8.1.1 Données

On choisit dans ce projet de s'intéresser au CAC40, on choisit de prendre les prix de fermeture du CAC40 du 1er janvier 2020 jusqu'à aujourd'hui, le 3 juin 2025. On récupère ces données grâce à R à l'aide des commandes suivantes :

```
>library("quantmod")
>getSymbols("^FCHI", src = "yahoo")
[1] "FCHI"
>cac40 <- Cl(FCHI)[2020-01-01/]
>plot(cac40)
>lines(SMA(cac40, n = 20), col = "red")
```

On obtient alors le graphique suivant :



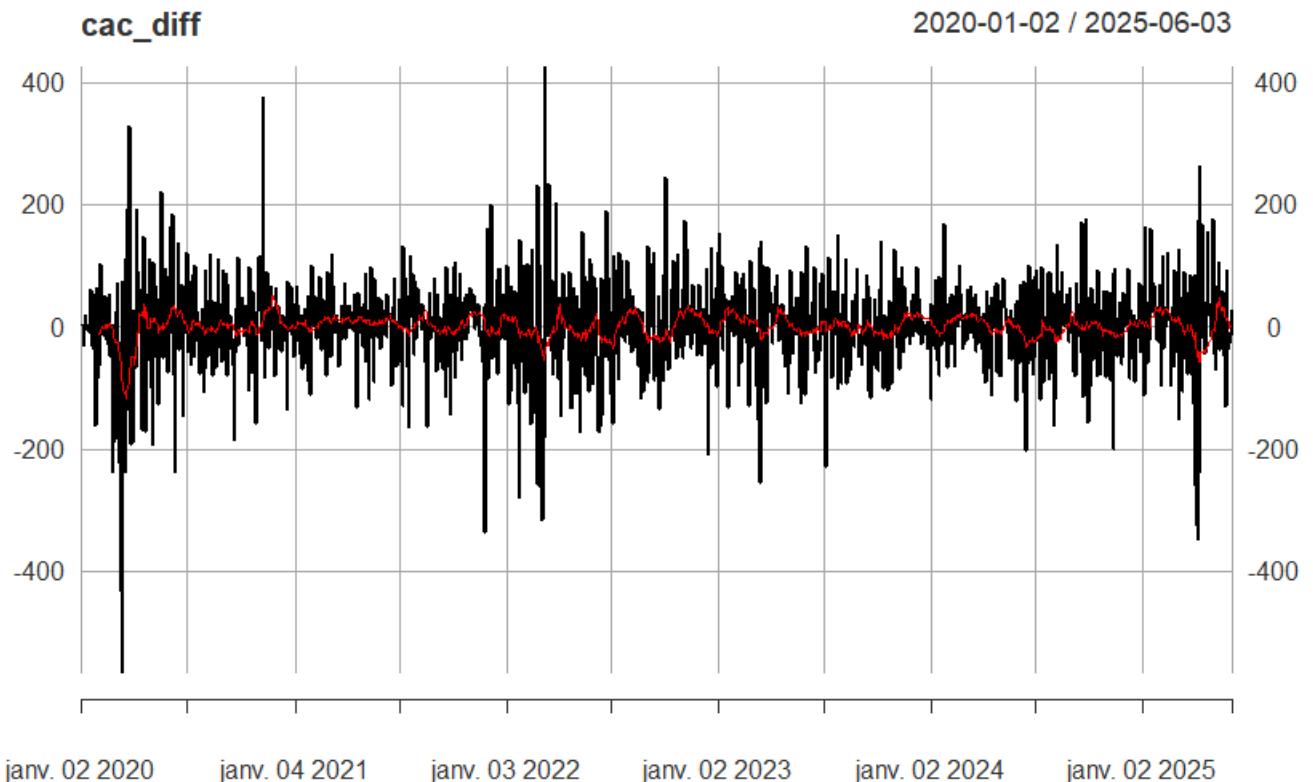
**Explication du graphique :** On observe sur ce graphique en noir, la séries temporelle donnant les prix de fermeture du CAC40 et en rouge la moyenne mobile sur 20 jours. Il est ici explicite visuellement que la série n'est pas stationnaire, elle a une tendance haussière.

#### 8.1.2 Stationnarité

Pour tenter de corriger cela, on transforme la série en soustrayant chaque valeur par la valeur précédente. Soit  $(Y_t)$  la série temporelle donnant les valeurs du CAC40, on crée la série temporelle  $(X_t)$  en la définissant comme :

$$X_t = Y_t - Y_{t-1} \iff X_t = (1 - L)^1 Y_t = \Delta^1(Y_t)$$

```
>cac_diff <- diff(cac40)
>plot(cac_diff)
>lines(SMA(cac_diff, n = 20), col = "red")
```



**Explication du graphique :** On remarque ici que la série semble suivre une moyenne, comme le montre la moyenne mobile en rouge. Il ne semble pas y avoir de tendance.

Nous réalisons le test de Dickey-Fuller pour vérifier que la série est bien stationnaire.

```
>adf.test(na.omit(cac_diff))

Augmented Dickey-Fuller Test

data : na.omit(cac_diff)
Dickey-Fuller = -10.398, Lag order = 11, p-value = 0.01
alternative hypothesis : stationary
```

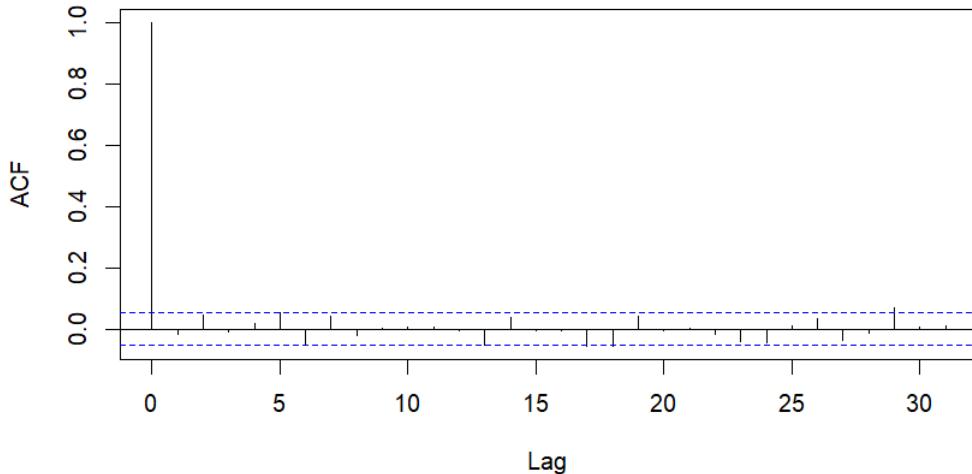
La p-value vaut  $p = 0.01 < 5\%$ , on en déduit que la série est stationnaire à 5%. On détermine donc  $d = 1$ , le modèle est maintenant ARIMA(p,1,q)

### 8.1.3 Paramètres p

Pour déterminer p et q du modèle d'ARIMA, on cherche à calculer la fonction d'autocorrélation et la fonction d'autocorrélation partielle.

```
>acf(na.omit(cac_diff), main = "Fonction d'autocorrélation")
```

### Fonction d'autocorrélation



La fonction d'autocorrélation chutte très rapidement ce qui confirme l'hypothèse de stationnarité de la série.

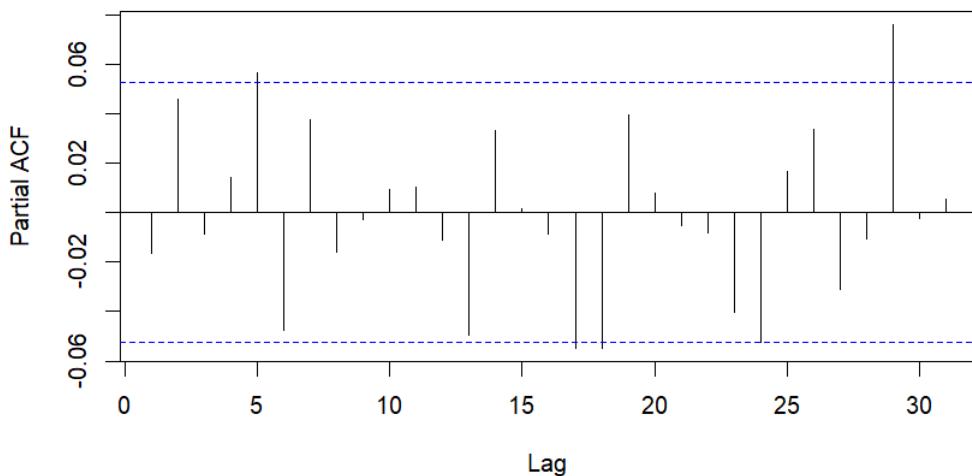
**Remarque :** On remarque que l'autocorrélation n'est significative qu'au rang 0, pour le reste, les autocorrélations sont insignifiantes. On suppose donc que  $q = 0$ .

Le modèle supposé est pour l'instant ARIMA(p,1,0).

#### 8.1.4 Paramètres q

```
>pacf(na.omit(cac_diff), main = "Fonction d'autocorrélation partielle")
```

### Fonction d'autocorrélation partielle



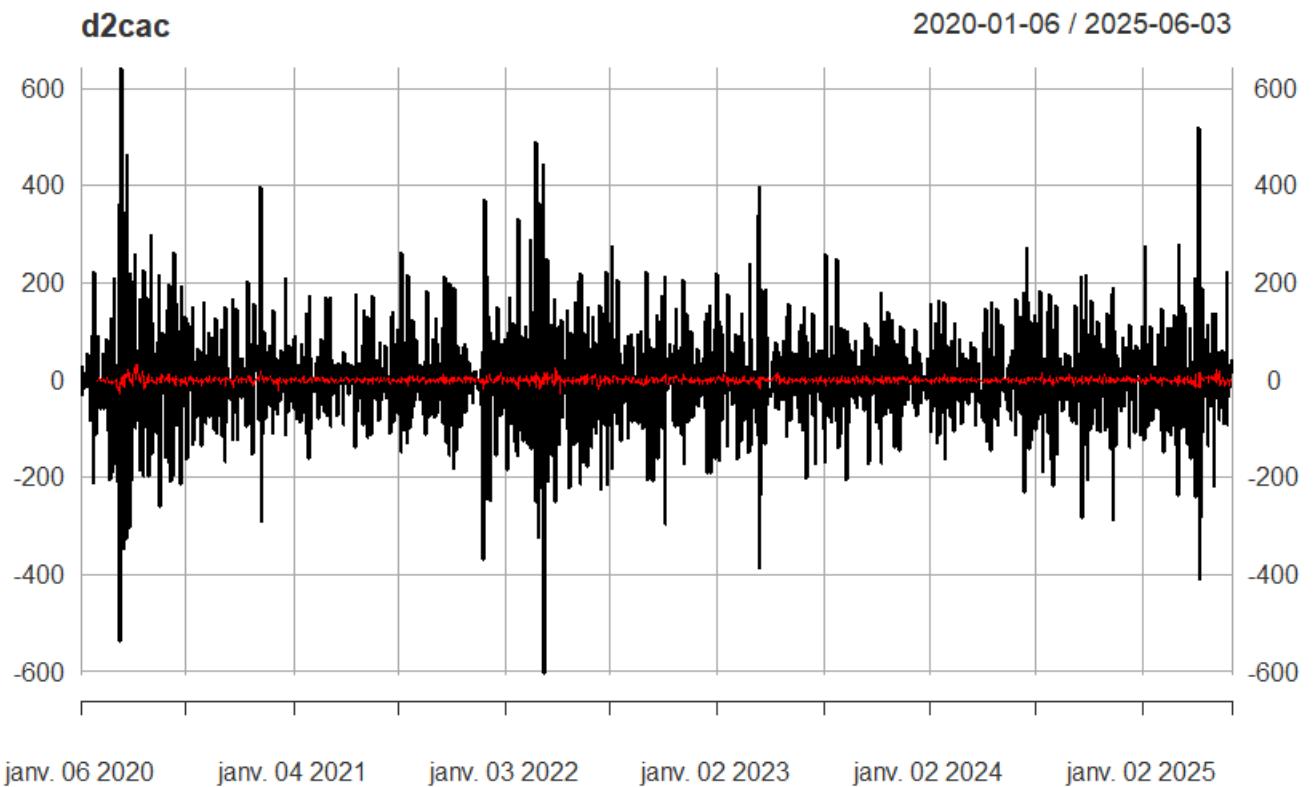
#### 8.1.5 Autocorrélation partielle trop significative

La fonction d'autocorrélation partielle est ici toujours significative, il est très compliqué de déterminer  $p$ . Nous allons différencier une deuxième fois la série afin de réduire l'autocorrélation partielle.

## 8.2 Identification avec $\Delta^2$

### 8.2.1 Stationnarité

```
>d2cac <- na.omit(diff(diff(cac40))
>plot(d2cac)
>lines(SMA(d2cac, n = 20), col = "red")
```



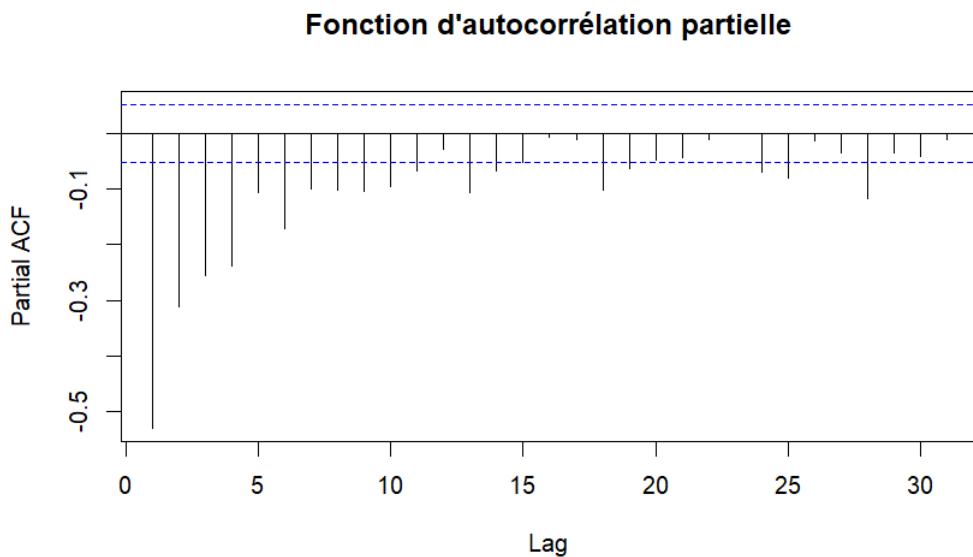
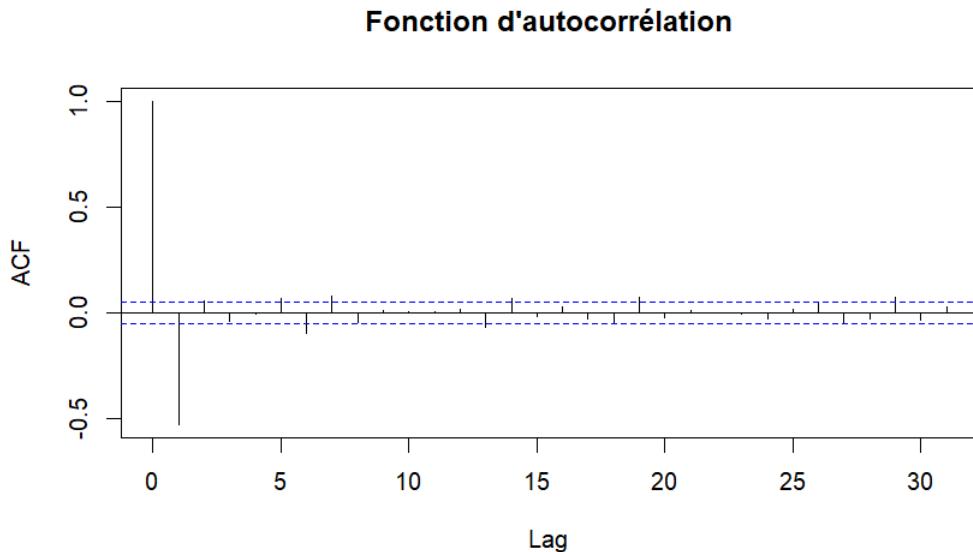
```
>adf.test(d2cac)

Augmented Dickey-Fuller Test

data : d2cac
Dickey-Fuller = -17.263, Lag order = 11, p-value = 0.01
alternative hypothesis : stationary
```

La série différenciée 2 fois est bien stationnaire. Le test de Dickey-Fuller et la moyenne mobile nous le montre.

### 8.2.2 Estimation de p et q



**Déduction des graphiques :** On voit que l'autocorrélation est significative jusqu'à ce que  $q = 1$ , l'autocorrélation partielle est quand à elle significative jusqu'à ce que  $p = 6$ .

**Choix du modèle :** Nous utiliserons donc le modèle ARIMA(6,2,1)

### 8.3 Estimations des paramètres

Nous avons donc le modèle ARIMA(6,2,1). L'équation suivante doit donc être vérifié :

$$\Phi(L)(1 - L)^2 Y_t = \Theta(L) \epsilon_t$$

où :

$$\begin{cases} \Phi(L) = \mathbb{I}_d - \phi_1 L - \phi_2 L^2 - \dots - \phi_6 L^6 \\ \Theta(L) = \mathbb{I}_d + \theta_1 L \end{cases}$$

donc :

$$(Y_t - 2Y_{t-1} + Y_{t+2})(\mathbb{I}_d - \sum_{i=1}^6 \phi_i L^i) = (\mathbb{I}_d + \theta_1 L) \epsilon_t$$

On pose :  $Z_t = Y_t - 2Y_{t-1} + Y_{t+2}$ , et on réécrit ainsi :

$$Y_t - 2Y_{t-1} + Y_{t+2} = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_6 Z_{t-6} + \theta_1 \epsilon_{t-1} + \epsilon_t$$

```
>install.packages("forecast")
>library(forecast)
>modele <- Arima(d2cac, order = c(6,2,1))
>summary(modele)
```

L'estimation des paramètres que R nous donne est l'estimation suivante :

TABLE 1 – Estimation des coefficients du modèle ARIMA(6,2,1)

Paramètre	Coefficient estimé	Erreur standard (s.e.)
$\phi_1$	-1.5254	0.0259
$\phi_2$	-1.6259	0.0454
$\phi_3$	-1.4612	0.0556
$\phi_4$	-1.1057	0.0555
$\phi_5$	-0.6350	0.0454
$\phi_6$	0.2669	0.0259
$\theta_1$	-1.0000	0.0020

On a donc le modèle suivant :

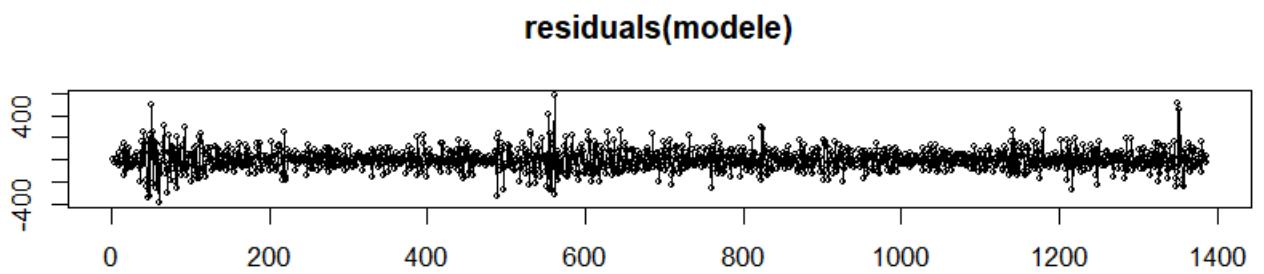
$$Z_t = -1.5254Z_{t-1} - 1.6259Z_{t-2} - 1.4612Z_{t-3} - 1.1057Z_{t-4} - 0.6350Z_{t-5} + 0.2669Z_{t-6} - 1.0000\epsilon_{t-1} + \epsilon_t$$

## 8.4 Vérification

### 8.4.1 Graphique

Nous allons représenter graphiquement les résidus afin de voir s'il semble bien être proche de 0.

```
>tsdisplay(residuals(modele))
```



### 8.4.2 Moyenne

```
>mean(residuals(modele))
[1] 0.0771026
```

Que ce soit graphiquement ou par une simple moyenne, la moyenne des résidus est très proche de 0. On a ici :  $\mathbb{E}[\epsilon_t] = 0.0771026 \approx 0$

## 8.5 Prédiction

Maintenant notre modèle estimé, on souhaite prédire les prochaines périodes afin d'estimer si le cours du CAC40 peut augmenter ou diminuer.

On rappelle que nous travaillons ici avec un modèle différencié deux fois. Autrement dit, notre modèle avec lequel nous travaillons n'est pas  $Y_t$  la valeur du CAC40 mais avec  $Z_t = Y_t - 2Y_{t-1} + Y_{t-2}$ .

Nous allons donc faire une prédiction mais directement sur  $Y_t$  afin de prédire les 15 prochaines périodes.

```
>modele2 <- Arima(cac40, order = c(6,2,1))
>pred <- forecast(modele2, h = 15)
>pred$mean
```

TABLE 2 – Estimation du CAC40 avec le modèle ARIMA(6,2,1)

Jour	Estimation
04/06/25	7762.050
05/06/25	7765.510
06/06/25	7764.838
09/06/25	7767.156
10/06/25	7770.387
11/06/25	7770.352
12/06/25	7771.921
13/06/25	7772.883
16/06/25	7774.341
17/06/25	7775.600
18/06/25	7776.693
19/06/25	7778.006
20/06/25	7779.211
23/06/25	7780.482
24/06/25	7781.707

## 8.6 Interprétation et Conclusion

La modélisation par le modèle d'ARIMA semble cohérente, en effet lorsque l'on observe le cours du CAC40, on voit que les variations d'un jour à l'autre sont plutôt croissantes, comme nous avons pu le voir sur le graphique de la série temporelle initiale.

Le modèle ARIMA est un bon modèle pour estimer la potentielle évolution d'une série temporelle. Il reste cependant incomplet car il ne prend pas en compte les événements politiques, économiques et financiers qui peuvent affecter le cours du CAC40.

## Bibliographie

- (1) température maximale à Melun <https://meteofrance.com/climat/releves/france/ile-de-france/melun>
- (2) Analyse des séries temporelles d'observations du système climatique, Mahdi Haddad, 2017 <file:///C:/Users/Arnaud/Downloads/analyse-des-s%C3%A9ries-temporelles-d%E2%80%99observations-du-syst%C3%A8me-climatique.pdf>
- (3) Séries Temporelles 2A, 2015 <https://ensai.fr/wp-content/uploads/2019/06/PolyseriesTemp.pdf>
- (4) Introduction aux séries temporelles, tendance et composante saisonnière, Yannig Goude [https://www.imo.universite-paris-saclay.fr/~yannig.goude/Materials/time\\_series/cours2\\_tendance\\_composante\\_saisonniere.pdf](https://www.imo.universite-paris-saclay.fr/~yannig.goude/Materials/time_series/cours2_tendance_composante_saisonniere.pdf)
- (5) COURS DE SERIES TEMPORELLES
- (6) THEORIE ET APPLICATIONS VOLUME 1, Arthur Charpentier <https://www.math.u-bordeaux.fr/~hzhang/m2/st/TS1.pdf>